TD 5 : Entropie

29 février 2024

1 Échauffement

Question 1 : On a une pièce biaisée de probabilité $0 < p < 1$ (on ne connaît pas p). Comment peut-on simuler une pièce parfaite en utilisant la pièce biaisée ? Combien de lancés sont-ils nécessaires en moyenne ?

Question 2 : On a une pièce parfaite ($p = 1/2$).

1. Comment peut-on simuler un dé à $n \in \{3, 4, 5, \dots\}$ faces ?
2. Quel est le temps moyen d'exécution ?
3. Peut-on réaliser cette simulation en temps borné ?

2 Entropie

Question 3 : Soit X une v.a. sur un domaine fini. Quelle est, en général, la relation d'inégalité entre $H(X)$ et $H(Y)$ si

a. $Y = 2^X$?

b. $Y = cos(X)$?

Question 4 : Quelle est la valeur minimale de $H(p_1, \dots, p_n) = H(\vec{p})$ pour $\vec{p}$ dans l'ensemble des vecteurs de probabilité de dimension n ? Donner tous les $\vec{p}$ qui réalisent ce minimum.

Question 5 : On lance une pièce parfaite jusqu'à obtenir face. La v.a. X dénote le nombre de lancers effectués.

a. Trouver $H(X)$.

b. On tire au hasard une v.a. X selon cette distribution. Trouver une séquence efficace de requêtes oui/non de la forme « Est-ce que X est dans l'ensemble S ? » pour déterminer X. Comparer $H(X)$ avec l'espérance du nombre de requêtes nécessaires pour trouver X.

Question 6 : Dans le cadre des InterENS Culturelles 2017, une série de 7 matchs d'improvisation est organisée entre les équipes de Paris-Saclay et de Rennes. La première équipe à remporter 4 matchs est déclarée gagnante. Soit X la v.a. représentant les résultats possibles, par exemple $SSSS$, $RSRSRSR$ ou $SSSRRRR$. Soit Y le nombre de matchs joués ($4 \leq Y \leq 7$). On suppose les équipes équilibrées, et les matchs indépendants. Calculer $H(X)$, $H(Y)$, $H(Y|X)$ et $H(X|Y)$.

3 Entropie dans les arbres

Définition 1 *Un arbre probabiliste est un arbre fini d'arité D dont les noeuds internes et les feuilles sont étiquetées par des probabilités, telles que*

(A) La probabilité de la racine vaut 1.

(B) La probabilité d'un noeud interne est la somme des probabilités de ses enfants.

On note :
- $P_1 \dots P_N$ les probabilités des N noeuds internes (P_1 étant la racine),
- $p_1 \dots p_n$ les probabilités des n feuilles,
- $q_{l,j}$ la probabilité du j-ième fils du l-ième noeud interne (ainsi $\forall l \; P_l = \sum_j q_{l,j}$)

Question 7 : Soit $F \in \{1 \dots n\}$ une variable aléatoire suivant la loi définie par les $p_1 \dots p_n$ et L la profondeur de la F-ième feuille. Montrer que

$$\mathbb{E}(L) = \sum_{i=1}^{N} P_i$$

Question 8 : Soit $W \in \{1 \dots N\}^*$ la suite de noeuds internes aboutissant à la feuille F et $B \in \{1 \dots D\}^*$ la suite de branchements correspondantes (ainsi $W[i+1]$ est le $B[i]$-ième enfant du noeud $W[i]$.) Soit l un noeud interne de profondeur i. Exprimer $H_l = H(B[i] = j \mid W[i] = l)$ en fonction des $q_{l,j}$ et de P_l.

Question 9 : Montrer que $H(F) = \sum_{l=1}^{N} P_l \cdot H_l$.

Question 10 : Soit un algorithme probabiliste effectuant L tirages indépendants de même distribution μ, et dont le résultat est la variable aléatoire Y. Montrer que $H(Y) \leq H(\mu) \times \mathbb{E}(L)$.